View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

JOURNAL OF ALGEBRA 146, 318–334 (1992)

On a Theorem of Ore

JESÚS MONTES

Departament d'Àlgebra i Geometria,
Facultat de Matemàtiques, Universitat de Barcelona,
Gran Via de les Corts Catalanes, 585, 08007 Barcelona, Catalunya, Spain

AND

ENRIC NART*

Departament de Matemàtiques,
Universitat Autònoma de Barcelona,
08193 Bellaterra, Barcelona, Catalunya, Spain

Communicated by A. Fröhlich

Received June 7, 1989

O. Ore (*Math. Ann.* **99**, 1928, 84–117) developed a method for obtaining the absolute discriminant and the prime-ideal decomposition of the rational primes in a number field K. The method, based on Newton's polygon techniques, worked only when certain polynomials $f_S(Y)$, attached to any side S of the polygon, had no multiple factors. These results are generalized in this paper finding a much weaker condition, effectively computable, under which it is still possible to give a complete answer to the above questions. The multiplicities of the irreducible factors of the polynomials $f_S(Y)$ play then an essential role. © 1992 Academic Press, Inc.

INTRODUCTION

Let K be a number field, A_K the ring of integers of K, and $\mathrm{disc}(K)$ its absolute discriminant. The determination of the prime-ideal decomposition in A_K of the rational primes, the computation of $\mathrm{disc}(K)$, and the construction of a basis of A_K as a $\mathbf{Z}$-module are three intimately related classical problems in algebraic number theory. After the work of Hensel, the three questions can be reduced to the local case, but even then, an effective solution in terms of a defining equation for K can be given only in an algorithmic way (cf. [1]). Among direct procedures to solve these problems, let

* Work on this paper has been partially supported by a Grant from C.A.I.C.Y.T. PB85-0075.

0021-8693/92 $3.00
Copyright © 1992 by Academic Press, Inc.
All rights of reproduction in any form reserved.

us mention a very partial answer given by a celebrated theorem of Dedekind, which is a refinement of a previous result of Kummer.

Let $f(X) \in \mathbf{Z}[X]$ be a monic irreducible polynomial, θ a root of $f(X)$, and $K = \mathbf{Q}(\theta)$. Let us denote by $\mathrm{disc}(f)$ the discriminant of $f(X)$ and by $\mathrm{ind}(f) = (A_K : \mathbf{Z}[\theta])$ the index of $f(X)$, so that

$$\mathrm{disc}(f) = \mathrm{ind}(f)^2 \cdot \mathrm{disc}(K).$$

Let $p \in \mathbf{Z}$ be a prime number and let $\bar{f}(X)$ be the polynomial of $\mathbf{F}_p[X]$ obtained by reducing the coefficients of $f(X)$ modulo p. Let

$$\bar{f}(X) = \varphi_1(X)^{e_1} \cdots \varphi_r(X)^{e_r} \tag{1}$$

be its factorization into a product of powers of distinct irreducible polynomials of $\mathbf{F}_p[X]$.

THEOREM OF KUMMER. *Suppose that* $p \nmid \mathrm{disc}(f)$, *that is,* $e_1 = \cdots = e_r = 1$. *Then,*

$$pA_K = \mathfrak{p}_1 \cdots \mathfrak{p}_r,$$

where, for all i, $\mathfrak{p}_i = pA_K + \varphi_i(\theta) A_K$ *is a prime ideal of* A_K *lying over* p *with residual degree* $f(\mathfrak{p}_i/p) = \deg \varphi_i(X)$.

THEOREM OF DEDEKIND. *Suppose that* $p \nmid \mathrm{ind}(f)$. *Then,*

$$pA_K = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r},$$

where, for all i, $\mathfrak{p}_i = pA_K + \varphi_i(\theta) A_K$ *is a prime ideal of* A_K *lying over* p *with ramification index* $e(\mathfrak{p}_i/p) = e_i$ *and residual degree* $f(\mathfrak{p}_i/p) = \deg \varphi_i(X)$.

In order to apply this last theorem in an effective way one needs a criterion to decide when the condition $p \nmid \mathrm{ind}(f)$ is satisfied. This was also supplied by Dedekind:

CRITERION OF DEDEKIND. *With the above notation, let* $g_1(X), \ldots, g_r(X)$ *be arbitrary monic polynomials of* $\mathbf{Z}[X]$ *such that* $\bar{g}_i(X) = \varphi_i(X)$ *and let*

$$g(X) = \frac{1}{p}\left(f(X) - g_1(X)^{e_1} \cdots g_r(X)^{e_r}\right).$$

Then, $p \nmid \mathrm{ind}(f)$ *if and only if for all* i *we have either* $e_i = 1$ *or* $\varphi_i(X) \nmid \bar{g}(X)$ *in* $\mathbf{F}_p[X]$.

Not as well known is a theorem of Ore which constitutes a far reaching

generalization of these results. By Hensel's lemma, we obtain from (1) a factorization in $\mathbf{Z}_p[X]$:

$$f(X) = f_1(X) \cdots f_r(X), \qquad f_i(X) = \varphi_i(X)^{e_i}.$$

If $p \mid \mathrm{ind}(f)$ these factors need not be irreducible in $\mathbf{Z}_p[X]$ and the problem is to determine its further decomposition into a product of irreducible factors. And even then, one has still to find, for each irreducible factor, the ramification index, residual degree, discriminant, and basis of integers of the corresponding local extension. Ore considers in [2] a Newton polygon for each i, whose sides provide a factorization of $f_i(X)$ in $\mathbf{Z}_p[X]$ and partial information about ramification indices, etc. Moreover, he attaches to each side S a polynomial $(f_i)_S(Y) \in \mathbf{F}_{q_i}[Y]$, $q_i = p^{\deg \varphi_i(X)}$, whose factorization in $\mathbf{F}_{q_i}[Y]$ provides a further factorization of the factor of $f_i(X)$ corresponding to this side. Finally, Ore shows that when all these polynomials $(f_i)_S(Y)$ have no multiple factors, then all these factors of $f_i(X)$ are irreducible, and the shape of the polygon and the degrees of the irreducible factors of the $(f_i)_S(Y)$ provide the necessary data to achieve a complete knowledge of all ramification indices and residual degrees (cf. Section 1).

The aim of this paper is to prove a generalization of these results of Ore in the same spirit as the Theorem of Dedekind generalizes that of Kummer. We find a condition, playing the same role as "$p \nmid \mathrm{ind}(f)$," under which it is still possible to obtain the complete decomposition of pA_K even when the polynomials $(f_i)_S(Y)$ have multiple factors. The multiplicities of these factors contribute then to the ramification indices of the prime ideals above p, as was to be expected. Finally, we find also an effective criterion, analogous to that of Dedekind, to decide when this condition is satisfied.

The method of Ore provides also an explicit formula for $\mathrm{ind}(f)$ (thus, an effective computation of $\mathrm{disc}(K)$) and allows one to construct a basis of the integers in a straightforward way. The same results are still valid in the more general situation that we consider.

We think that our result can be a key step to develop a very fast algorithm for obtaining prime ideal decomposition and integral basis. The major virtue of the algorithm will be that only polynomial-factoring routines over finite *fields* are needed.

1. The Work of Ore on Newton Polygons

In this section we give a short review of the paper of Ore [2] and we introduce some concepts and notation to be used in the rest of the paper without further mention.

We fix a prime number $p \in \mathbf{Z}$ and algebraic closures $\bar{\mathbf{Q}}_p$, $\bar{\mathbf{F}}_p$ of $\mathbf{Q}_p$ and $\mathbf{F}_p$. For any finite extension L of $\mathbf{Q}_p$ we shall denote by A_L the ring of integers, $\mathfrak{p}_L$ the prime ideal of L, v_L the standard valuation of $\bar{\mathbf{Q}}_p$ suitably normalized to satisfy $v_L(L^*) = \mathbf{Z}$, and by $\mathbf{F}_L$ the residue field, which we shall assume embedded in $\bar{\mathbf{F}}_p$ in such a way that all reduction maps $A_L \to \mathbf{F}_L$, denoted $a \mapsto \bar{a}$, become naturally compatible. For any monic irreducible polynomial $\psi(X) \in \mathbf{F}_L[X]$ we fix once and for all a monic polynomial in $A_L[X]$ reducing to $\psi(X)$ modulo $\mathfrak{p}_L$, and we denote it by the same symbol $\psi(X)$.

We take a finite extension K of $\mathbf{Q}_p$ as a ground field and denote $A = A_K$, $\mathfrak{p} = \mathfrak{p}_K$, $v = v_K$, and $\mathbf{F} = \mathbf{F}_K$. We fix throughout a prime element $\pi \in A$, a monic irreducible polynomial $\varphi(X) \in \mathbf{F}[X]$ of degree $m \geqslant 1$ and a root $\zeta \in \bar{\mathbf{F}}_p$ of $\varphi(X)$. Let T be the non-ramified extension of K of degree m. We have $\mathbf{F}_T = \mathbf{F}(\zeta)$.

Let $f(X) \in A[X]$ be a polynomial of degree $n \geqslant 1$. It can be written in only one way as

$$f(X) = \sum_{i=0}^{[n/m]} a_i(X) \cdot \varphi(X)^i, \tag{2}$$

with $a_i(X) \in A[X]$ and $\deg a_i(X) < m$ or $a_i(X) = 0$. Let s_i be the greatest exponent such that π^{s_i} divides all the coefficients of $a_i(X)$. The $\varphi(X)$-polygon of $f(X)$ is the lower convex envelope of the set of points (i, s_i) in the Euclidean plane. The classical Newton polygon corresponds to the case $\varphi(X) = X$. The typical shape of this polygon in the case of a monic polynomial is shown in Fig. 1. The set of sides with negative slope constitutes the "principal part" of the polygon. Its projection onto the X-axis has length equal to the greatest exponent l such that $\varphi(X)^l$ divides $f(X)$ in $\mathbf{F}[X]$.

We come now to one of the main inventions of Ore, which we give in a slightly more general setting, as we shall need it in the next sections. Let S be any segment in the Euclidean plane such that (r, s) and $(r + E, s - H)$,

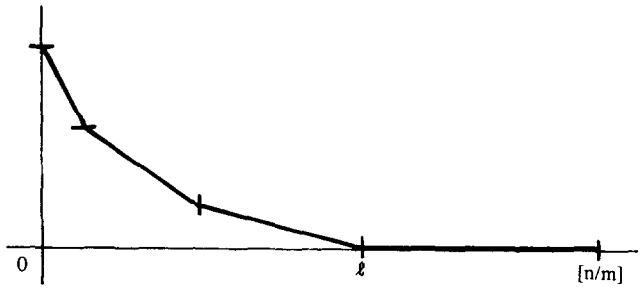

FIGURE 1

r, s, E, H non-negative integers, are the first and last points with integer coordinates belonging to S. Let us define

$$d = HCF(E, H), \qquad e = E/d, \qquad h = H/d, \tag{3}$$

with the convention that $e = h = 0$ when $E = H = 0$. Suppose that in the vertical lines passing through the points of integer coordinates belonging to S there are no points of the $\varphi(X)$-polygon of $f(X)$ lying below S. Then we can associate to each of these points an element of $\mathbf{F}_T$ and take these elements as the coefficients of a polynomial $f_S(Y) \in \mathbf{F}_T[Y]$ which will be called simply the *polynomial associated to* $f(X)$ *and* S. In fact, the polynomials

$$b_j(X) = \pi^{-s+jh} a_{r+je}(X), \qquad 0 \leqslant j \leqslant d,$$

have integer coefficients and we can define

$$f_S(Y) = \sum_{j=0}^{d} \bar{b}_j(\zeta) \cdot Y^j \in \mathbf{F}_T[Y].$$

The non-zero coefficients of $f_S(Y)$ correspond to points of the $\varphi(X)$-polygon of $f(X)$ belonging to S. In particular, in the case considered by Ore, that S is precisely one of the sides of the $\varphi(X)$-polygon of $f(X)$, this polynomial $f_S(Y)$ will have degree exactly d and non-zero constant term. For simplicity we shall always assume in this case that $f_S(Y)$ denotes the monic polynomial of $\mathbf{F}_T[Y]$ obtained by dividing by the principal term.

We can group the main results of [2] in four theorems. The first two are typical of Newton polygons techniques:

THEOREM OF THE PRODUCT. *Let* $f_1(X), ..., f_r(X) \in A[X]$ *be monic polynomials and take* $f(X) = f_1(X) \cdots f_r(X)$. *The principal part of the* $\varphi(X)$*-polygon of* $f(X)$ *is made by joining all sides of the principal parts of the* $\varphi(X)$*-polygons of all* $f_i(X)$ *in descreasing order of slopes. Moreover, for each side* S *with negative slope of the* $\varphi(X)$*-polygon of* $f(X)$ *we have*

$$f_S(Y) = (f_{i_1})_{S_1}(Y) \cdots (f_{i_t})_{S_t}(Y),$$

where $\{i_1, ..., i_t\} \subseteq \{1, 2, ..., r\}$ *and* $S_1, ..., S_t$ *are all the sides with the same slope as* S *belonging to any of the* $\varphi(X)$*-polygons of* $f_1(X), ..., f_r(X)$.

THEOREM OF THE POLYGON. *Let* $f(X) \in A[X]$ *be a monic polynomial and suppose that the* $\varphi(X)$*-polygon of* $f(X)$ *has* r *different sides* $S_1, ..., S_r$ *with slopes* $-h_1/e_1 < \cdots < -h_r/e_r$. *Then* $f(X)$ *admits a factorization,* $f(X) = f_1(X) \cdots f_r(X)$, *where each* $f_i(X)$ *is a monic polynomial in* $A[X]$ *whose* $\varphi(X)$*-polygon consists of only one side with the same shape as* S_i *and such*

that the polynomial of $\mathbf{F}_T[Y]$ *associated to this side is equal to* $f_{S_i}(Y)$. *Moreover, for all* i, *if* θ *is a root of* $f_i(X)$ *we have* $v(\varphi(\theta)) = h_i/e_i$.

Remark. The concept of $\varphi(X)$-polygon and these two results add nothing new to the classical theory. In fact, assume (just for simplicity) that all sides of the $\varphi(X)$-polygon of $f(X)$ have negative slope, that is, that $\bar{f}(X) = \varphi(X)^l$. We can find a polynomial $g(X) \in T[X]$ such that

$$f(X) = \prod_{\sigma \in \mathrm{Gal}(T/K)} g^{\sigma}(X) \qquad \text{and} \qquad \bar{g}(X) = (X - \zeta)^l.$$

Since T/K is unramified, it is easy to see that each of the questions we are interested in about $f(X)$ can be reduced to the same question about $g(X)$, taking T as the ground field. To see this, denote also by ζ the root of $\varphi(X)$ in $\bar{\mathbf{Q}}_p$ reducing to ζ modulo $\mathfrak{p}_T$, make the linear change of variables $g_1(X) = g(X + \zeta)$, and consider the classical Newton polygon of $g_1(X)$. The reader will check easily that this polygon has exactly the same shape as the $\varphi(X)$-polygon of $f(X)$ and that for each side, the associated polynomial is the same in both polygons after the change of variables $y \mapsto \varphi'(\zeta) \cdot Y$.

Nevertheless, we shall maintain throughout the rest of the paper the language of $\varphi(X)$-polygons in the statements of the theorems, since it has the advantage of displaying the results in a completely effective way in terms of the original polynomial we are interested in. However, it is clear from what we have just remarked that we can restrict ourselves in the proofs to the case of classical polygons.

The really interesting contribution of Ore lies in the next theorem, where he shows that a factorization of $f_{S_i}(Y)$ into coprime factors in $\mathbf{F}_T[Y]$ ensures a further factorization of the polynomial $f_i(X)$.

THEOREM OF ORE. *Let* $f(X) \in A[X]$ *be a monic polynomial with one-sided* $\varphi(X)$*-polygon* S *and let* e, h *be defined as in* (3). *Let*

$$f_S(Y) = \psi_1(Y)^{e_1} \cdots \psi_t(Y)^{e_t}$$

be the factorization of $f_S(Y)$ *into a product of powers of distinct irreducible polynomials of* $\mathbf{F}_T[Y]$. *Then* $f(X)$ *admits a factorization,* $f(X) = f_1(X) \cdots f_t(X)$, *where each* $f_i(X)$ *is a monic polynomial of* $A[X]$ *with one-sided* $\varphi(X)$*-polygon* S_i *of the same slope as* S *and associated polynomial* $(f_i)_{S_i}(Y) = \psi_i(Y)^{e_i}$.

Moreover, if $e_1 = \cdots = e_t = 1$, *then* $f_1(X), \ldots, f_t(X)$ *are already irreducible and, for all* i, *if* θ *is a root of* $f_i(X)$ *and* $L = K(\theta)$ *we have* $\mathfrak{p}_L = (\varphi(\theta)^b/\pi^c)\mathfrak{A}_L$, *where* b, c *are positive integers such that* $bh - ce = 1$, *and*

$$e(L/K) = e, \qquad f(L/K) = m \cdot \deg \psi_i(Y).$$

This theorem is completely satisfactory from the point of view of effectiveness. If we deal with an arbitrary polynomial $f(X)$, the theorem is meant to be applied to some factor of $f(X)$ corresponding to one side of some polygon of $f(X)$, but the only data we need to know about this factor is the shape and the polynomial associated to this side, and both invariants can be read directly from the polygon of the original polynomial.

Also important and very much necessary for our purposes is a formula of Ore for the index of $f(X)$. We need some more notation: let $f(X) \in A[X]$ be a monic irreducible polynomial, θ a root of $f(X)$, and $L = K(\theta)$. We have

$$(A_L : A[\theta]) = p^s, \qquad s \geqslant 0,$$

and we define

$$i(f) = i_K(f) = v(p^s)/[K: \mathbf{Q}_p] = s/f(\mathfrak{p}/p).$$

This value is always an integer and we still have the relation

$$v(\mathrm{disc}(f)) = v(\mathrm{disc}(L/K)) + 2i(f).$$

If $f(X)$ is any monic polynomial of $A[X]$ and $f(X) = \prod_i f_i(X)$ is its decomposition as a product of monic irreducible polynomials of $A[X]$, we define

$$i(f) := \sum_i i(f_i) + \sum_{i<j} r(f_i, f_j), \tag{4}$$

where $r(f_i, f_j)$ denotes the $\mathfrak{p}$-valuation of the resultant of $f_i(X)$ and $f_j(X)$. Note that this relation (4) holds for any decomposition of $f(X)$, even when the factors are not necessarily irreducible.

Assume now that $f(X)$ is a monic irreducible polynomial of $\mathbf{Z}[X]$. The p-exponent of the usual index of $f(X)$ deserves also this notation: $i(f) = v_p(\mathrm{ind}(f))$. We get in this way two meanings for $i(f)$ when we think of $f(X)$ belonging to $\mathbf{Z}[X]$ or to $\mathbf{Z}_p[X]$. The definition (4) has been introduced in order to make them coincide.

For each irreducible polynomial $\varphi(X) \in \mathbf{F}[X]$ and monic polynomial $f(X) \in A[X]$ we shall denote by $i_\varphi(f)$ deg $\varphi(X)$ times the number of points of integer coordinates below or belonging to the $\varphi(X)$-polygon of $f(X)$, excluding those lying on both axis; that is,

$$i_\varphi(f) = \deg \varphi(X) \left(\sum_{i=1}^{r-1} E_i \cdot \left(\sum_{j=i+1}^{r} H_j \right) + \frac{1}{2} \sum_{i=1}^{r} (H_i E_i - H_i - E_i + d_i) \right),$$

where $S_1, \ldots, S_r$ are the sides of the principal part of the $\varphi(X)$-polygon of

$f(X)$ and, for each side, the parameters E_i, H_i, d_i are defined as in (3). Clearly $\varphi(X) \nmid \bar{f}(X)$ in $\mathbf{F}[X]$ implies that $i_\varphi(f) = 0$, but the converse is obviously not true.

THEOREM OF THE INDEX. *Let $f(X) \in A[X]$ be a monic polynomial. Then,*

$$i(f) \geqslant \sum_{\varphi(X)} i_\varphi(f),$$

as $\varphi(X)$ runs over the monic irreducible polynomials of $\mathbf{F}[X]$ dividing $\bar{f}(X)$. Moreover, if all the polynomials associated to all sides of all $\varphi(X)$-polygons have no multiple factors then equality holds.

2. THE MAIN THEOREM

If $f(X) \in A[X]$ has one-sided $\varphi(X)$-polygon S, by the Theorem of the Index we have $i(f) \geqslant i_\varphi(f)$, and $f_S(Y)$ having no multiple factors is a sufficient condition that ensures the equality. Precisely this property, $i(f) = i_\varphi(f)$ is the weaker condition to play the role analogous to $i(f) = 0$ in the Theorem of Dedekind.

THEOREM 1. *Let $f(X) \in A[X]$ be a monic polynomial with one-sided $\varphi(X)$-polygon S and let e, h be defined as in* (3). *Let*

$$f_S(Y) = \psi_1(Y)^{e_1} \cdots \psi_t(Y)^{e_t}$$

be the factorization of $f_S(Y)$ into a product of powers of distinct irreducible polynomials of $\mathbf{F}_T[Y]$. Then $f(X)$ admits a factorization, $f(X) = f_1(X) \cdots f_t(X)$, where each $f_i(X)$ is a monic polynomial of $A[X]$ with one-sided $\varphi(X)$-polygon S_i of the same slope as S and associated polynomial $(f_i)_{S_i}(Y) = \psi_i(Y)^{e_i}$. Moreover, if $i(f) = i_\varphi(f)$, then all $f_i(X)$ are irreducible and if θ is a root of $f_i(X)$ and $L = K(\theta)$ we have

$$\mathfrak{p}_L = (\varphi(\theta)^b/\pi^c) A_L \qquad \text{or} \qquad \psi_i(\varphi(\theta)^e/\pi^h) A_L,$$

according as $e_i = 1$ or $e_i > 1$, where b, c are positive integers such that $bh - ce = 1$, and

$$e(L/K) = e \cdot e_i, \qquad f(L/K) = m \cdot \deg \psi_i(Y).$$

As far as effectiveness is concerned, this theorem is useless without a criterion to decide when the condition $i(f) = i_\varphi(f)$ is satisfied. At this point, motivated by the Criterion of Dedekind, we proceed as follows. For

each i let $\psi_i(X, Y) \in A[X, Y]$ be any preimage of $\psi_i(Y) \in \mathbf{F}_T[Y]$ under the canonical homomorphism $A[X, Y] \to \mathbf{F}[X, Y]/\varphi(X) \cong \mathbf{F}_T[Y]$, satisfying

$$\psi_i(X, Y) = \sum_j a_j(X)\, Y^j, \qquad \deg a_j(X) < m \text{ or } a_j(X) = 0.$$

Let H be the length of the projection of S onto the Y-axis and define polynomials

$$f^0(X) = \pi^H \cdot \prod_i \psi_i(X, \varphi(X)^e/\pi^h)^{e_i} \in A[X],$$

and $f^1(X) = f(X) - f^0(X)$. All points of the $\varphi(X)$-polygon of $f^1(X)$ lie strictly above S and the points at a minimal distance from S contain the information we are interested in. These points belong to the side $\tilde{S}$ parallel to S obtained by shifting S upwards until it touches a point with integer coordinates (see Fig. 3). The object playing a role analogous to that of the polynomial $\bar{g}(X)$ in Dedekind's Criterion is the polynomial $f^1_{\tilde{S}}(Y)$ associated to $f^1(X)$, $\varphi(X)$, and $\tilde{S}$:

CRITERION 1. *With the above notation, $i(f) = i_\varphi(f)$ if and only if for each i we have either $e_i = 1$ or $\psi_i(Y) \nmid f^1_{\tilde{S}}(Y)$ in $\mathbf{F}_T[Y]$.*

Remark. In the most common case that the slope of S is not an integer (that is, when $e > 1$), this polynomial $f^1_{\tilde{S}}(Y)$ is particularly easy to obtain. The points of the $\varphi(X)$-polygon of $f^1(X)$ belonging to $\tilde{S}$ are exactly those of the polygon of $f(X)$ previously lying there. In fact, the points of the $\varphi(X)$-polygon of $f(X)$ belonging to S, resp. to $\tilde{S}$, have abscissa divisible by e, resp. congruent to j modulo e, where $0 < j < e$ is the solution to $jh \equiv 1 \pmod{e}$; hence, they all belong to different vertical lines. When we take the difference with $f^0(X)$, the ordinates of the former points are increased by at least one whereas the latter remain unchanged. Thus, we can read $f^1_{\tilde{S}}(Y)$ directly from the polyfon of $f(X)$ without computing $f^0(X)$ and $f^1(X)$. *In other words, in this case we have $f^1_{\tilde{S}}(Y) = f_{\tilde{S}}(Y)$.*

For the proof of Theorem 1 and Criterion 1 we need a couple of lemmas and one (crucial) proposition:

LEMMA 1 (Ore). *Let $f(X) \in A[X]$ be a monic irreducible polynomial such that $\bar{f}(X) = \varphi(X)^E$. Let S be the side of the $\varphi(X)$-polygon of $f(X)$ and let e, h be the geometrical data of S given by (3). Let θ be a root of $f(X)$ such that $\bar{\theta} = \zeta$ and let $\psi(Y)$ be the minimal polynomial of $\varphi(\theta)^e/\pi^h$ over $\mathbf{F}_T$. Then, $f_S(Y)$ is a power of $\psi(Y)$.*

Proof. As we remarked in the preceding section we can reduce to the

case $\varphi(X) = X$. Let us denote the element θ^e/π^h by γ and let $g(Y)$ be the minimal polynomial of γ over K. Let

$$g_1(X) = \pi^{h \cdot \deg g(Y)} \cdot g(X^e/\pi^h).$$

We have $v(\gamma) = 0$, so that $g(Y)$ and $g_1(X)$ have integer coefficients and the constant term of $g(Y)$ has valuation zero. Thus, the polygons of $g(Y)$ and $g_1(X)$ are one-sided, of slope zero and $-h/e$, respectively. The polynomial associated to $g_1(X)$ and this side of slope $-h/e$ is precisely $\bar{g}(Y)$, which is a power of $\psi(Y)$. On the other hand, $g_1(\theta) = 0$ implies that $f(X)$ divides $g_1(X)$ and by the Theorem of the Product, $f_S(Y)$ is also a power of $\psi(Y)$. ∎

LEMMA 2. *Let $f(X)$, $g(X) \in A[X]$ be monic polynomials of degree n, n', respectively. Suppose that their $\varphi(X)$-polygons consist of only one side S, S' with projection onto the Y-axis of length H, H', respectively. Then, the* p*-valuation of the resultant of $f(X)$ and $g(X)$, $r(f, g) = v(\mathrm{Res}(f, g))$, satisfies*

$$r(f, g) \geqslant \min\{nH', n'H\},$$

and equality holds if and only if either the slopes of S and S' are different or $f_S(Y)$ and $g_{S'}(Y)$ have no common factor.

Proof. Let E, H, d, e, h; E', H', e', h' be the respective geometrical data of S and S' given by (3) and suppose that $H'/E' \geqslant H/E$. If the $\varphi(X)$-expansion of $f(X)$ is given by (2) we define

$$f_0(X) = \sum_{j=0}^{d} a_{je}(X) \cdot \varphi(X)^{je}.$$

For any root ω of $g(X)$ we have $v(\varphi(\omega)) = H'/E' \geqslant H/E$, so that $v(f(\omega)) \geqslant H$, and $v(f(\omega)) > H$ if and only if $v(f_0(\omega)) > H$; this last condition is equivalent to

$$v(\varphi(\omega)) = H/E \qquad \text{and} \qquad f_S^{\sigma}(\overline{\varphi(\omega)^e/\pi^h}) = 0,$$

where $\sigma \in \mathrm{Gal}(T/K)$ is the unique automorphism such that $\sigma(\zeta) = \bar{\omega}$. But $H'/E' = H/E$ implies that $e = e'$ and $h = h'$, so that by Lemma 1 we have that $\overline{\varphi(\omega)^e/\pi^h}$ is a root of $g_{S'}^{\sigma}(Y)$, precisely for the same automorphism σ. Therefore, the condition $v(f(\omega)) > H$ is equivalent to

$$H'/E' = H/E \qquad \text{and} \qquad f_S(Y),\ g_{S'}(Y) \text{ have a common root.}$$

Since $r(f, g)$ is the sum of the $v(f(\omega))$ for all roots ω of $g(X)$, the lemma is proven. ∎

In the next proposition we prove the theorem under the assumption that $f(X)$ is irreducible. In this case, by the Theorem of the Polygon and Lemma 1, the $\varphi(X)$-polygon has only one side S and $f_S(Y)$ is a power of an irreducible polynomial.

PROPOSITION. *Let $f(X)\in A[X]$ be a monic irreducible polynomial of degree $n\geqslant 1$. Then $i(f)\geqslant i_\varphi(f)$. Moreover, let S be the only side of the $\varphi(X)$-polygon of $f(X)$, let E, H, d, e, h be the geometrical data given by* (3) *and let $\psi(Y)\in \mathbf{F}_T[Y]$ be the monic irreducible polynomial such that $f_S(Y)=\psi(Y)^a$, $a\geqslant 1$. Let θ be a root of $f(X)$, $L=K(\theta)$, and*

$$\gamma_i=\varphi(\theta)^i/\pi^{[iH/E]}\in A_L,\qquad i\in\mathbf{Z},\ i\geqslant 0.$$

Then, the following conditions are equivalent:

(i) $i(f)=i_\varphi(f)$.

(ii) $\gamma_0,\ldots,\gamma_{E-1}$ *are a basis of A_L as A_T-module.*

(iii) $e(L/K)=e\cdot a$, $f(L/K)=m\cdot\deg\psi(Y)$, *and either* $a=1$ *or* $\mathfrak{p}_L=\psi(\gamma_e)A_L$.

(iv) *Either* $a=1$ *or* $\psi(Y)\nmid f_S^1(Y)$.

(v) *Either* $a=1$ *or* $A_L=A_T[\gamma_e]$.

Proof. Clearly $K\subseteq T\subseteq L$ and $f(X)=\prod_{\sigma\in \mathrm{Gal}(T/K)} g^\sigma(X)$, where $g(X)\in A_T[X]$ is the minimal polynomial of θ over T. Since T/K is unramified we have $i_K(f)=m\cdot i_T(g)$ and it is easy to check, by the Remark of Section 2, that we can reduce to the case $\varphi(X)=X$. We have then $E=n=[L:K]$.

Let A_0 be the sub-A-module (in fact an order) of A_L generated by $\gamma_0,\ldots,\gamma_{n-1}$. It is clear that

$$v((A_0:A[\theta]))=[K:\mathbf{Q}_p]\cdot\sum_{i=0}^{n-1}[iH/n],$$

and for each i, $[iH/n]$ is the number of points with integer coordinates below or belonging to S, with abscissa $n-i$. This proves the first assertion of the Proposition and proves also that the conditions (i) and (ii) are equivalent.

The following general properties of the γ_i are clear:

$$\begin{gathered} v(\gamma_i)=0 \qquad \text{if and only if } e\mid i.\\ \gamma_{je}=\gamma_e^j \qquad \text{for all } \ 0\leqslant j<d.\\ \{v(\gamma_1),\ldots,v(\gamma_{e-1})\}=\{1/e,2/e,\ldots,(e-1)/e\}.\\ \gamma_i=\gamma_r\cdot\gamma_e^j, \qquad \text{if } \ i=je+r,\ 0\leqslant r<e. \end{gathered}\tag{5}$$

We shall end the proof of the Proposition by showing that (ii) $\Rightarrow$ (iii) $\Leftrightarrow$ (iv) and (iii) $\Rightarrow$ (v) $\Rightarrow$ (ii). For convenience we shall denote $\gamma = \gamma_e$ and $t = d/a = \deg \psi(Y)$.

Assume that $\gamma_0, ..., \gamma_{n-1}$ constitute an A-basis of A_L. Their images under the reduction map must contain an $\mathbf{F}$-basis of $\mathbf{F}_L$, but the only elements with non-zero reduction are γ^j, $0 \leqslant j < d$, and among their reductions, $\{\bar{\gamma}^j,\ 0 \leqslant j < t\}$ is a maximal subset of $\mathbf{F}$-independent elements, since they are the powers of a root of the irreducible polynomial $\psi(Y) \in \mathbf{F}[Y]$. Therefore $f(L/K) = t$ and, in consequence, $e(L/K) = ae$. We have still to show that either $a = 1$ or $v_L(\psi(\gamma)) = 1$. Let $0 < i_0 < e$ be the subindex such that $v(\gamma_{i_0}) = (e-1)/e$. The element $\omega = \gamma_{i_0} \cdot \psi(\gamma)^{a-1} \in A_L$ is equal to $\gamma_{i_0} \cdot \gamma^{(a-1)\cdot t}$ plus a sum of terms of the type $u \cdot \gamma_{i_0} \cdot \gamma^b$, $u \in A$, $b < (a-1)t$. Since $i_0 + e(a-1)t$ is less than n, this is already the expression of ω as the A-linear combination of the basis $\gamma_0, ..., \gamma_{n-1}$. Thus, $v_L(\omega) < e(L/K)$ since not all the coefficients of this expression are divisible by π. Now,

$$v_L(\omega) = (e-1)a + (a-1)\, v_L(\psi(\gamma)) < ea,$$

leads to $a > (a-1)v_L(\psi(\gamma))$, and this implies that either $a = 1$ or $v_L(\psi(\gamma)) = 1$. We have thus shown that (ii) $\Rightarrow$ (iii).

Since $e(L/K)$ is always a multiple of e (by the Theorem of the Polygon) and $f(L/K)$ is always a multiple of t, when $a = 1$ we have already $e(L/K) = e$ and $f(L/K) = t$. Assume now that $a > 1$. From $0 = f(\theta) = f^0(\theta) + f^1(\theta)$ we get $v(f^0(\theta)) = v(f^1(\theta))$; hence, we have the following chain of equivalences:

$$\begin{aligned}(\text{iii}) &\Leftrightarrow v(\psi(\gamma)) = 1/ea \Leftrightarrow v(f^0(\theta)) = H + (1/e) \\ &\Leftrightarrow v(f^1(\theta)) = H + (1/e) \Leftrightarrow f_S^1(\bar{\gamma}) \neq 0 \Leftrightarrow (\text{iv}).\end{aligned}$$

The last but one equivalence can be obtained by arguing as in the proof of Lemma 2 and the last equivalence is a consequence of Lemma 1.

If $f(L/K) = t$, then $\{\bar{\gamma}^j,\ 0 \leqslant j < t\}$ constitutes an $\mathbf{F}$-basis of $\mathbf{F}_L$. If, moreover, $v_L(\psi(\gamma)) = 1$ then $\{\gamma^j \cdot \psi(\gamma)^i,\ 0 \leqslant j < t,\ 0 \leqslant i < ae\}$ is an A-basis of A_L, so that $A_L \subseteq A[\gamma]$. Thus (iii) $\Rightarrow$ (v).

Let us show now that A_0 is a subring of A_L. Since $\gamma_i \cdot \gamma_j$ is either γ_{i+j} or $\pi\gamma_{i+j}$, it is sufficient to show that γ_i belongs to A_0 for all $i \geqslant 0$. If $i \geqslant n$, it is clear that γ_i belongs to the A-module generated by $\gamma_{i-1}, ..., \gamma_{i-n}$ and by induction, to A_0. We have to prove that (v) $\Rightarrow$ (ii). If $a = 1$ we have $A_L = A_0$ by (5), since $e(L/K) = e$ and $f(L/K) = d$. Finally, since A_0 is a subring of A_L we have always $A_L \supseteq A_0 \supseteq A[\gamma]$. This ends the proof of the Proposition. ∎

Proof of Theorem 1. The first part of Theorem 1, which is common to the Theorem of Ore, is an immediate consequence of Lemma 1 and the

Theorem of the Product. Assume now that the condition $i(f) = i_\varphi(f)$ is satisfied.

By the Theorem of the Product, S is made up of pieces S_i, each one being the side of the $\varphi(X)$-polygon of $f_i(X)$. Let E_i, H_i be the length of the respective projections of S_i onto the X-axis and Y-axis. By Lemma 2 we have $r(f_i, f_j) = (\deg f_i) \cdot H_j = m \cdot E_i \cdot H_j$ (recall that $m = \deg \varphi(X)$), and this value is equal to m times the number of points of integer coordinates in the rectangle determined by the projections of S_i, S_j onto the X-axis and Y-axis, respectively (see Fig. 2). The whole set of points below S can be distributed in triangles and rectangles as shown in Fig. 2. An analogous distribution, for each i, of the points in the triangle corresponding to S_i, shows, by the Proposition and Lemma 2, that $i(f_i)$ is greater than or equal to the number of points of integer coordinates in that triangle. Therefore, the condition $i(f) = i_\varphi(f)$ is equivalent to $i(f_i) = i_\varphi(f_i)$ for all i. Thus, the theorem will follow from the Proposition if we show that all the factors $f_i(X)$ are irreducible. Let $f_i(X) = g(X) \cdot h(X)$ be any factorization of $f_i(X)$ as a product of two monic polynomials of $A[X]$ of degree $\geqslant 1$. Let $R = i_\varphi(f_i) - i_\varphi(g) - i_\varphi(h)$. We have just seen that $i(g) \geqslant i_\varphi(g)$ and $i(h) \geqslant i_\varphi(h)$. But, by Lemma 2, we have $r(f, g) > R$ since the polynomials of $\mathbf{F}_T[Y]$ associated to the respective sides of the $\varphi(X)$-polygons of $g(X)$ and $h(X)$ are both powers of $\psi_i(Y)$. Thus, $i(f_i) > i_\varphi(f_i)$, in contradiction with our hypothesis. ∎

Proof of Criterion 1. As we saw in the proof of Theorem 1, the condition $i(f) = i_\varphi(f)$ is equivalent to $i(f_i) = i_\varphi(f_i)$ for all i, and this is equivalent by the Proposition to either $e_i = 1$ or $\psi_i(Y) \nmid (f_i)_{S_i}^1$ for all i. Since $f^0(X) = \prod_i f_i^0(X)$, we have

$$f^1(X) = \sum_i f_i^1(X) \cdot \prod_{j \neq i} f_j^0(X) + h(X), \tag{6}$$

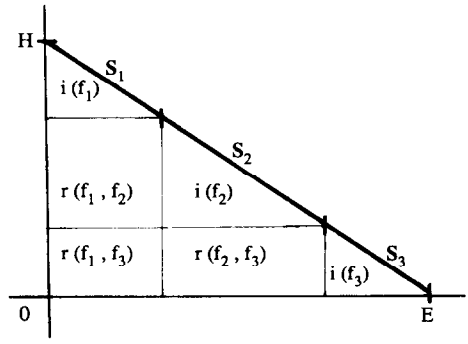

FIGURE 2

and the points of the $\varphi(X)$-polygon of $h(X)$ lie strictly above $\tilde{S}$. Since all the sides S_i have the same slope it is easy to check that

$$f_{\tilde{S}}^{1}(Y) = \sum_i (f_i)_{\tilde{S}_i}^{1}(Y) \cdot \prod_{j \neq i} (f_j)_{S_j}(Y).$$

Therefore, for all i, the condition $\psi_i(Y) \mid (f_i)_{\tilde{S}_i}^{1}(Y)$ is equivalent to $\psi_i(Y) \mid f_{\tilde{S}}^{1}(Y)$. ∎

We finish this section with an example illustrating the practical application of Theorem 1 and Criterion 1. Let $f(X) \in \mathbf{Z}[X]$ be any monic irreducible polynomial having classical Newton polygon with the shape indicated in Fig. 3, that is,

$$f(X) = X^6 + p(a_5 X^5 + a_4 X^4) + p^2(a_3 X^3 + a_2 X^2) + p^3(a_1 X + a_0),$$

with $a_i \in \mathbf{Z}$ such that $p \nmid a_0$. Let θ be a root of $f(X)$ and $K = \mathbf{Q}(\theta)$. There are six points of integer coordinates below S, hence $i(f) = v_p(\mathrm{ind}(f)) \geqslant 6$. The associated polynomial is

$$f_S(Y) = Y^3 + \bar{a}_4 Y^2 + \bar{a}_2 Y + \bar{a}_0 \in \mathbf{F}_p[Y].$$

Assume that Ore's condition is not satisfied: there exist $a, b \in \mathbf{Z}$ such that either

$$f_S(Y) = (Y - \bar{a})^3, \qquad \text{or} \qquad f_S(Y) = (Y - \bar{a})^2 (Y - \bar{b}).$$

Criterion 1 says that $i(f) = 6$ if and only if $\bar{a}$ is not a root of the polynomial

$$f_{\tilde{S}}^{1}(Y) = f_{\tilde{S}}(Y) = \bar{a}_5 Y^2 + \bar{a}_3 Y + \bar{a}_1.$$

In this case, Theorem 1 tells us that

$$pA_K = \mathfrak{p}^6, \qquad \mathfrak{p} = \left(\frac{\theta^2}{p} - a\right) A_K + pA_K,$$

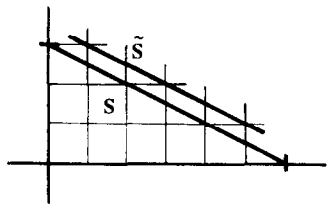

FIGURE 3

or

$$pA_K = \mathfrak{p}^4(\mathfrak{p}')^2, \qquad \mathfrak{p} = \left(\frac{\theta^2}{p} - a\right) A_K + pA_K, \mathfrak{p}' = \frac{\theta^3}{p} A_K + pA_K,$$

respectively. Moreover,

$$\left\{1, \theta, \frac{\theta^2}{p}, \frac{\theta^3}{p}, \frac{\theta^4}{p^2}, \frac{\theta^5}{p^2}\right\}$$

is an integral basis of the p-localization of A_K.

3. Many-Sided Polygons

Let $f(X) \in A[X]$ be an arbitrary monic polynomial, let $S_1, ..., S_r$ be the sides of the $\varphi(X)$-polygon of $f(X)$, and let $f_1(X), ..., f_r(X) \in A[X]$ be the corresponding factors of $f(X)$ given by the Theorem of the Polygon. Let us denote one of these pairs $f_i(X)$, S_i generically by $g(X)$, S. If we want to obtain the arithmetical data about $f(X)$ that we are interested in throughout this paper, we have to be able to apply Theorem 1 and Criterion 1 to $g(X)$ is an effective way, but without having to compute $g(X)$ explicitly. The data we need to know is the shape of S and the two polynomials $g_S(Y)$, $g_S^1(Y)$. The Theorem of the Polygon shows that the first two invariants can be obtained directly from the polygon of $f(X)$, but, what about $g_S^1(Y)$?. We have only to find an effective criterion to decide which irreducible factors of $g_S(Y)$ divide $g_S^1(Y)$. The most natural procedure is to consider the polynomial

$$f^1(X) = f(X) - \prod_i f_i^0(X),$$

and hope that the points of the $\varphi(X)$-polygon of $f^1(X)$ will give us the desired information. It is as natural as this, except for some perverse subtlety. In fact, the polynomial $f_S^1(Y)$ is not only different from $g_S^1(Y)$, but even fails to be divisible by the same irreducible factors of $g_S(Y)$. Thus, we cannot apply Criterion 1 with this polynomial. This mishap is easy to explain: we have the same relation (6) and again all points of the $\varphi(X)$-polygon of $h(X)$ lie strictly above $\tilde{S}$, so that this polynomial adds no interesting information and we can ignore it. If $S = S_i$, we can read off $g_S^1(Y)$ from the points furnished by the summand of (6) corresponding to the subindex i. The trouble comes from the fact that some terms of the other summands can also furnish points belonging to $\tilde{S}$, which added to the others can disturb the information previously obtained. In order to

control this phenomenon we have to enlarge the side $\tilde{S}$ until it cuts the two straight lines determined by the adjacent sides $\tilde{S}_{i-1}$, $\tilde{S}_{i+1}$. If we denote by $\bar{S}$ this enlarged side we have:

CRITERION 2. *With the above notations, for any irreducible factor* $\psi(Y)$ *of* $g_S(Y)$ *we have* $\psi(Y) \mid g^1_{\bar{S}}(Y)$ *if and only if* $\psi(Y) \mid f^1_{\bar{S}}(Y)$.

Proof. It is tedious to explain it in detail but straightforward to check it. The point is that only the summands of (6) corresponding to the subindices $i-1$ and/or $i+1$ can furnish points of the $\varphi(X)$-polygon belonging to $\tilde{S}$, and this can happen only if $\bar{S}$ meets $\bar{S}_{i-1}$ and/or $\bar{S}_{i+1}$ in a point with integer coordinates. Even in this case, the two/one disturbing summands of (6) contribute to $f^1_{\bar{S}}(Y)$ with a multiple of $f_S(Y) = g_S(Y)$. On the other hand, the enlargement of $\tilde{S}$ can affect the contribution of the i-summand to $f^1_{\bar{S}}(Y)$ only by multiplying it by some power of Y. Thus, in any case the polynomial $f^1_{\bar{S}}(Y)$ is of the type

$$f^1_{\bar{S}}(Y) = c \cdot Y^s \cdot g^1_{\tilde{S}}(Y) + h(Y) \cdot g_S(Y),$$

where $h(Y) \in \mathbf{F}_T[Y]$, $c \in \mathbf{F}_T$, $c \neq 0$, and $s \geqslant 0$. Since the irreducible factors of $g_S(Y)$ are always different from Y, the criterion is proven. ∎

If the slope of S is $-h/e$ and the slope of the side to the right (left) is $-h'/e'$, it is easy to see that it is necessary to enlarge $\tilde{S}$ to the right (left) if and only if $h'e - he'$ divides $(h-h', e-e')$ and $e < e'$. These sides $\bar{S}$ can be very much longer than $\tilde{S}$ or even reduce to a point. In Fig. 4 we given an example showing both possibilities.

The reader will now be able to combine Lemmas 1 and 2, the Proposition, and the geometrical arguments of the proof of Theorem 1 to prove the following strengthening of the Theorem of the Index:

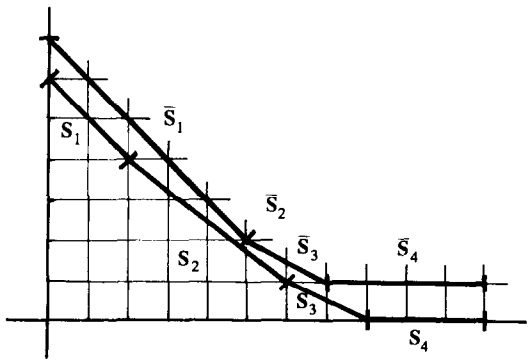

FIGURE 4

THEOREM 2. *Let $f(X) \in A[X]$ be a monic polynomial. Then,*

$$i(f) \geqslant \sum_{\varphi(X)} i_{\varphi}(f),$$

as $\varphi(X)$ runs over the monic irreducible polynomials of $\mathbf{F}[X]$ dividing $\bar{f}(X)$. Moreover, the equality holds if and only if for any side S of any $\varphi(X)$-polygon and for any irreducible factor $\psi(Y)$ of $f_S(Y)$ we have either $\psi(Y)^2 \nmid f_S(Y)$ or $\psi(Y) \nmid f_S^1(Y)$.

REFERENCES

1. R. BÖFFGEN, Der Algorithmus von Ford/Zassenhaus zur Berechnung von Ganzheitsbasen in Polynomalgebren, *Ann. Univ. Sarav.* **1**, No. 3 (1987).
2. Ö. ORE, Newtonsche Polygone in der Theorie der algebraischen Körper, *Math. Ann.* **99** (1928), 84–117.